I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

On the Wisdom of Categorical (Privacy) Rules

KIEL BRENNAN-MARQUEZ

CONTENTS

I. INTRODUCTION

Privacy often collides with other values,[1] forcing the law to adopt strategies to manage the tradeoff. This essay explores those strategies. In it, I have three goals.

First, I distinguish between two mutually-exclusive ways of managing the tradeoff between privacy and other values: "case-by-case balancing," which asks whether, all things considered, a particular violation of privacy is justified by countervailing benefits, and the "categorical approach," which asks whether the party seeking

[1] Law enforcement and national security are the most obvious examples and will be the focus here. Others include: scientific research (especially in medicine); commercial innovation, see, e.g., Adrienne LaFrance, *Privacy Problems with Driverless Cars*, THE ATLANTIC (Mar. 24, 2016), http://www.theatlantic.com/technology/archive/2016/03/self-driving-car-makers-on-privacy-just-trust-us/474903/ [https://perma.cc/75UK-LVZC]; and certain forms of social accountability that involve the publication of otherwise private information, like journalism. In fact, it turns out that privacy even trades off against itself. *See, e.g.*, David Pozen, *Privacy-Privacy Tradeoffs*, 83 U. CHI. L. REV. 221 (2016).

permission to violate privacy has met a generally-applicable burden, like probable cause. Second, I show that Helen Nissenbaum's "contextual integrity" framework,[2] the basis of her and Paula Kift's luminous analysis in *Metadata in Context*,[3] is an example of the former. Third, I defend the categorical approach on normative grounds, and thereby cast doubt on the utility of contextual integrity analysis for surveillance law. There is great wisdom, I argue, in preventing courts (and law enforcement authorities) from indulging in case-specific cost-benefit analyses, and instead tethering privacy law to stable *ex ante* benchmarks.

II. CASE-BY-CASE BALANCING V. CATEGORICAL RULES

When gains in one dimension cause losses in another,[4] one strategy for managing the tradeoff is to ask whether the gains outweigh the losses in a particular case. Another strategy is to fashion categorical decision-rules by which to resolve particular cases *without* recourse to a case-specific analysis of gains and losses. Such decision-rules can certainly be informed by cost-benefit analysis ("CBA") — indeed, it would be odd if CBA did *not* play some role in the formulation of decision-rules — but once the rules are set, they have a preemptive effect on CBA, case-by-case.

Existing constitutional and statutory rules regulating state surveillance exemplify the categorical approach. To determine the legality of a given search (or seizure), the question is *not* whether the search's benefits outweigh its costs. The question is whether law enforcement satisfied a specific evidentiary burden before it carried out the search — typically, probable cause.[5] If law enforcement failed

---

[2] For background, see HELEN NISSENBAUM, PRIVACY IN CONTEXT (2010).

[3] Paula Kift & Helen Nissenbaum, *Metadata in Context: An Ontological and Normative Analysis of the NSA's Bulk Telephony Metadata Collection Program*, 13 I/S: J.L. & POL'Y FOR INFO. SOC'Y 333 (2017).

[4] For simplicity's sake, I am using "two dimensions" as shorthand; the same analysis would apply, *mutatis mutandis*, to multi-variant normative problems.

[5] To be sure, not all searches require probable cause. But the differences between probable cause and other evidentiary burdens, such as reasonable suspicion, exigency, or relevance, are immaterial to my analysis. The point is that *all* evidentiary burdens are "categorical," in the sense that I am using the term. For helpful background, see CHRISTOPHER SLOBOGIN, PRIVACY AT RISK 21-49 (2007). Probable cause thus operates a synecdoche for the categorical approach *writ large*.

to satisfy that burden, the search or seizure is unlawful.[6] Period. In other words, if the police (say) enter a private home without probable cause, we prohibit the government from arguing that the benefits of entering the home were great enough, in this particular instance, to justify overriding the general rule. Whether we are *wise* to prohibit this style of argument is the focus of Part III. For now, the point is simply to appreciate that the prohibition exists, and more to the point, that the reason behind the prohibition is not that the premise of the government's argument (i.e., that the benefits of intrusion can outweigh the costs of intrusion in a particular case) is wrong. Rather, the reason for the prohibition is that *even assuming the government's premise is right*, we have decided that it will not carry the day; indeed, that it is irrelevant to the legal analysis.

III. KIFT AND NISSENBAUM'S ANALYSIS

With that in mind, what should we make of Kift and Nissenbaum's claim that bulk metadata collection — such as that carried out by the NSA, largely in secret, under Section 215 of the PATRIOT Act after 9/11 — offends privacy under a contextual integrity ("CI") analysis? Surely they are right about the offense. It would be difficult to argue with a straight face that bulk data collection programs, "meta" or otherwise, raise no privacy concerns.[7] Indeed, one of richest parts of their article is how it unearths the precarious, perhaps even incoherent, foundation of the supposed distinction between "data" and "metadata." For this accomplishment alone, the article will stand as a lasting contribution.

The natural next question, however, is whether (and in what sense) Kift and Nissenbaum's analysis can help reform the legal rules that constrain state surveillance. On this front, I am skeptical. The CI framework, as Kift and Nissenbaum describe it, is a cost-benefit framework. Indeed, it is a highly refined one. The central insight of CI is that when people transmit information to others, the significance of

---

[6] For analytic purposes, I am treating exceptions to a burden as part of the burden, insofar as exceptions operate as legal justifications. *See* Malcolm Thornbun, *Justifications, Power, and Authority*, 117 YALE L.J. 1070, 1103-07 (2008) (explaining that warrants operate as justifications, in the same manner as affirmative defenses in the criminal law). In other words, if the police can satisfy one of the exceptions to the probable cause requirement (for example), clearly their conduct is lawful. But that is just to say that they have met a different burden. It does not make the analysis any less categorical.

[7] Though government lawyers manage it surprisingly often.

the transmission is context-bound. In particular, expectations about the subsequent flow of information, which Kift and Nissenbaum call "transmission principles" — whether the information will be kept confidential, whether it will be transmitted to another party, and so forth — are defined entirely by context. This provides an easy and useful way of talking about privacy harms: a privacy harm results from the violation of transmission principles. If A transmits information to B and, because of contextual norms, A expects the information to be kept confidential, B's decision to transmit the information to C will result in a privacy harm to A.

So far, so good. But that is not where the CI analysis ends. When information flow violates a transmission principle (in the example, B's transmission of information to C), "a prima facie case exists for claiming [a privacy harm]."[8] But this "prima facie assessment" does not necessarily condemn the errant flow.[9] Rather, the errant flow can be appropriate, notwithstanding the privacy harm it produces, if it better promotes the "values, goals and ends of a given context" than adherence to the transmission principle would have.[10]

With regard to Section 215, for example, Kift and Nissenbaum make it quite clear that in their view, the bulk metadata collection program was (1) a violation of transmission principles (specifically, those governing information flow from users to telecom companies), and more importantly (2) an inappropriate violation of transmission principles. Ultimately, however, the driving force behind this conclusion is an empirical claim about the program's *efficacy*. Although the metadata collection program clearly "constitute[d] a prima facie violation" of privacy, Kift and Nissenbaum concede that "theoretically," the program could "still be justified" under contextual integrity principles if it "ultimately better achieve[d] the goal of national security, namely the prevention of future terrorist attacks."[11] The problem, they argue, is that the program did not actually *do* this, despite the government's repeated claims that "the bulk telephony

---

[8] Kift & Nissenbaum, *supra* note 3, at 16.

[9] *Id.*

[10] *Id.*

[11] *Id.* at 29.

metadata collection program contributed to thwarting over 50 different terrorist attacks."[12]

I take no sides in this empirical dispute. My point is simply that Kift and Nissenbaum admit, as they must, that privacy collides with other values — in the case of Section 215, national security — and the strategy they adopt for managing that tradeoff, hewing to CI principles, is to ask, in the context of a particular information flow, how the privacy harm measures up against countervailing benefits. As a matter of normative theory, this approach seems quite sensible. The question is whether it can hope to inform the regulation of state surveillance.

Regrettably, I think not. In this area, decision-rules are categorical; paradigmatically, searches and seizures must be based on probable cause.[13] If metadata collection qualifies as a search, as Kift and Nissenbaum suggest (joining a chorus of other scholars in the process),[14] probable cause, or something like it, should be required to justify specific instances of collection. Arguments about the national security outcomes enabled by collection, notwithstanding the *absence* of probable cause, should not prevail. In fact, they should not even get off the ground. For the same reason that, if the police barge into a private home without probable cause, the government would get no mileage by pointing to the search's salutary consequences, so the government should get no mileage by pointing to the salutary consequences of specific instances of metadata collection.

Naturally, this does not mean the (in some cases considerable) benefits of state surveillance have no role to play in legal analysis. It means, rather, that their role is to help *set* categorical decision-rules, not to override those rules case-by-case. In other words, if particular surveillance practices such as metadata collection carry enormous benefits, that may be a reason to authorize those practices more often.

---

[12] *Id.*

[13] Sometimes, decision rules have multiple components. For example, to enter a home under the Fourth Amendment's exigency exception to warrant requirement, law enforcement must be able to show exigency and probable cause. *See, e.g.*, Kentucky v. King, 131 S.Ct. 1849, 1856 (2011) (discussing the exigency rule). Needless to say, the multiplication of categorical rules does not make the resulting composite rule any less categorical.

[14] The literature along these lines is voluminous. For a particularly useful and comprehensive take, see Laura Donohue, *Bulk Metadata Collection: Statutory and Constitutional Considerations*, 37 HARV. J.L. & PUB. POL'Y 757 (2014).

This could be achieved, for example, by relaxing the standard that state officials must satisfy as a precondition of surveillance authority.[15]

But there are two important caveats here. First, the fact that a surveillance practice serves worthy goals is a necessary but insufficient condition for relaxing the state's evidentiary burden. The other side of the coin is intrusiveness. Suppose we had good evidence suggesting that if police cameras were running 24 hours a day in every home across the United States, it would wipe out 99% of violent crime. Even so, I suspect that very few of us would entertain the idea of permitting such a program. Efficacy aside, the privacy costs are simply too high. We would still demand probable cause.

Second, even if the goals advanced by surveillance (balanced against its intrusiveness) justify imposing a lesser evidentiary burden on the state, the resulting burden would still operate categorically. There is still no room, in other words, for case-specific override. This makes it difficult to see how Kift and Nissenbaum's consequentialist gambit — the argument that information flows, even if they violate transmission principles, can nonetheless "be justified" if they "promote[] the values, goals and ends of a given context" — translates into law. Traditionally, the rules governing state surveillance do not focus on global justifiability; they focus instead on the satisfaction of specific evidentiary burdens. Some information flows qualify as searches, and are therefore subject to the probable cause requirement, *regardless* of how well or poorly they "promote the values, goals [or] ends" of the context in which they occur. The hypothetical from a moment ago — the stationing of police cameras inside private homes — is an example. Likewise, to my mind, is bulk metadata collection.

IV. THE WISDOM OF CATEGORICAL (PRIVACY) RULES

Traditions, of course, are not normative arguments. In response to the categorical approach just described, one could easily argue that privacy law *ought* to abide by case-by-case balancing — that privacy law's faith in categorical decision-rules is misplaced — especially when the interests on the other side, such as thwarting terrorism, are of such manifest importance.

---

[15] This could happen in a variety of ways. For example, downgrading the evidentiary burden from "probable cause" to "reasonable suspicion," or even replacing suspicion with an adequately-enforced "relevance" requirement (including, e.g., provisions to protect against abusive or pretextual searches).

This Part offers a brief defense of categorical privacy rules. In broad strokes, I argue that categorical rules draw strength from three principles. The first concerns the nature of case-by-case adjudication: there are limits to the CBA that we can expect police officers and judges to perform. The second principle concerns the nature of privacy harms, which often seem minimal in the context of particular cases, but become recognizably significant in the aggregate. This principle counsels in favor, I argue, of addressing fact-patterns by category rather than in isolation. And last but not least, the third principle concerns the value of predictable legal rules.

A. *The Practical Limits of Cost-Benefit Analysis*

The first tick in favor of categorical privacy rules is the difficulty that would arise from actually *doing* CBA case-by-case. To enrich a hypothetical from earlier, suppose the police barge into a private residence, concededly without probable cause, and once inside they find a missing child ("Billy") whom they have been pursuing for the last two days. When the search is challenged, Lt. Smith, the supervising officer, offers the following explanation: "I had a strong hunch that Billy was in the house. We did not have enough for a warrant, but I just got the sense that something was wrong. So I decided to go in. I know the standard is probable cause — and beyond that, a warrant. But I thought a departure was justified under the circumstances, given the high stakes."

Under CBA, this argument may win the day, or it may not. We would need many more facts — a full picture of the totality of circumstances — to evaluate it thoroughly. Even in the abstract, however, it is easy see how vexing it will be to fashion evaluative criteria. The question at first appears straightforward: why not just balance the benefit associated with the finding the lost child against the cost to the homeowner's privacy? But on reflection, its complexity quickly multiplies.

The complexity has many facets, but I would worry, specifically, about two. First, we would need to determine the scope of relevant costs. Should the measure of costs in this particular case include, for example, the likelihood that if Lt. Smith's justification is accepted, the police will be marginally more likely to flout the probable cause and warrant requirements in the next case? If this cost is relevant — and I see no plausible argument why it would not be — a full interrogation of that question (an entire trial?) would be necessary to perform coherent CBA as to the specific intrusion. Second, we would need to

ask whether adherence to the categorical rule in fact *would not* have secured the relevant benefit. In other words, if the police had (1) performed further investigation, (2) drudged up facts sufficient to yield probable cause to search the house where Billy was found, and (3) secured a warrant, would they still have found Billy? This question is hard to answer. In fact, it is hard to know where to begin. The problem is similar to that of "harmless error" analysis in the due process context. Questions that turn on counterfactual imagination, requiring one to explore the infinitude of things that might have happened, as opposed to what actually happened, pose inherent epistemic difficulties.

These issues could certainly be explored further; philosophers have been writing about them for centuries. But I trust the point is clear. Whether because we do not trust judges or police officers to perform the foregoing interpretive tasks correctly, or simply because we think the costs of performing them are simply too high to bear, we wisely dispense with case-specific inquiry. Who, after all, would favor a standard for authorizing police intrusion that requires a full analysis of systemic effects in each case?

B. *Harm-Aggregation*

The second tick in favor of categorical privacy rules is that they facilitate the aggregation of privacy injuries, which is often necessary to grasp the full extent of harm. Even bracketing the conceptual difficulties of case-by-case analysis, in particular cases (like poor Billy's) there is often an imbalance between privacy harms, on one hand, and countervailing benefits, on the other, leading to a serious potential for distortive effects.

Take the hypothetical explored above. Lt. Smith's team enters a home without probable cause based on a hunch that Billy is captive inside. Suppose we stipulate that entering the home without normal cause was necessary to find Billy: had the police waited for a warrant, Billy would not have been rescued. Suppose we further stipulate that Billy's case is sufficiently idiosyncratic that permitting the intrusion on cost-benefit grounds is unlikely to make the police significantly more likely to flout Fourth Amendment rules going forward.

Measuring cost against benefit is still no easy task. If the question is framed as balancing (1) the harm of unwarranted entrance into the residence against (2) the benefit of rescuing Billy, the answer becomes self-evident: Billy should be rescued. But what if the question is framed, instead, as balancing (1) the harm of living in a world where

police intrude into homes without cause, whenever they have sufficiently strong hunches of wrongdoing, against (2) the benefits, in the aggregate, that such intrusions stand to yield?

I am not saying the second question would necessarily come out differently than the first; that is an empirical question. The point, rather, is that although both questions aim at the same issue — does the harm of intrusion outweigh its benefits, or vice versa? — they invite two rather different inquiries. To ask the question in aggregated form is just *different* from asking it in isolated form. In aggregated form, the question prompts us to consider not only the harm wrought about by intrusion for each particular homeowner, but also the collective harm of living in a society where police are empowered to make judgment calls about when (and when not) to follow constitutional rules. That harm matters. Constitutional law is not simply about redressing one-off grievances. It is also about creating a world in which we go about our daily lives free from concern about unchecked government intrusion. The aggregation of privacy harms is crucial to that task.

C. *The Virtues of Stability*

The final tick in favor of categorical privacy rules is not specific to privacy law; it is a virtue of categorical decision-rules in general. Simply put, they promote stability. They allow us to plan our lives.[16] Thus far in this Part, I have been talking about what it would mean to let police argue around the probable cause standard and warrant requirement by pointing out the benefits of a specific intrusion. But the argument would also run the other way. On the same grounds that police could argue for a suspension of the warrant requirement, victims of intrusion could also presumably argue that *adherence to* the warrant requirement is *insufficient* to justify an intrusion. In other words, if the question centers on case-by-case CBA, there is no *a priori* reason to think that compliance with a categorical rule would always strike the correct balance. On the contrary, CBA is an appealing analytic tool precisely because categorical rules *cannot* be trusted, across cases, to strike the right balance. What is more, case law is already rife with fact-patterns in which probable cause was

[16] *See* SCOTT SHAPIRO, LEGALITY (2011) (developing a "planning" theory of law — i.e., describing law as a mechanism for allocating interpretive authority to different actors in predictable ways).

satisfied —*Whren v. United States* and *Atwater v. City of Lago Vista*, perhaps most glaringly[17] — but the costs of intrusion seem, to many observers, to significantly outstrip its benefits.

Which raises the question: do we want a world in which probable cause (and other categorical rules) are demoted to the level of rebuttable presumptions? I think not. And especially not in areas like policing, where categorical rules, for better or worse, stabilize interactions between adversarial parties that would otherwise be prone to escalate. In this sense, categorical rules solve an important coordination problem.[18] They negotiate boundaries in advance and, in doing so, provide actors on the ground with concrete heuristics to guide decision-making. Absent such heuristics, the police would either become too zealous — capitalizing on the latitude afforded by consequentialist exceptions to probable cause — or not zealous enough — fearing for potential challenges even when probable cause exists — or, somewhat paradoxically, both at once.

V. CONCLUSION

In the coming decades, striking the proper balance between privacy and other values will be among the foremost challenges that judges, legislators, and regulators confront. In confronting it, they will have a choice to make: should the law preserve its categorical approach to privacy intrusions, or does the age of big data call for a more granular style of CBA?

In this Essay, I sought to highlight the virtues of the categorical approach. Naturally, this is not to say that cost-benefit frameworks — like Kift and Nissenbaum's — have no role to play in privacy law. But we should be clear about what that role is: to *refine* categorical decision-rules, not to supplant them.

---

[17] *See* Atwater v. City of Lago Vista, 532 U.S. 318 (2001) (holding that no Fourth Amendment violation occurred, due to the existence of probable cause, when the police arrested a woman for driving without a seatbelt and hauled her and her small children to the local precinct — an action that even the majority deemed "pointless"); Whren v. United States, 517 U.S. 806 (1996) (holding that probable cause is sufficient to justify a search, even if the search was carried out for pretextual, including race-motivated, reasons). *See also* Devenpeck v. Alford, 543 U.S. 146 (2004) (holding that an arrest was justified, due to the existence of probable cause to arrest for Crime X, despite the fact that the police understood themselves to be performing an arrest for Crime Y).

[18] For background on "coordination" as a concept from game theory, see Roger B. Myerson, *Justice, Institutions, and Multiple Equilibria*, 5 CHI. J. INT'L L. 91 (2004).